POT: estimators for the tail and the quantile of the excess distribution

POT: estimators for the tail and the quantile of the excess distribution

Let $X_1, \ldots, X_n$ i.i.d. r.v. with distribution function $F \in MDA(H_\gamma)$ for $\gamma \in \mathbb{R}$.

POT: estimators for the tail and the quantile of the excess distribution

Let $X_1, \ldots, X_n$ i.i.d. r.v. with distribution function $F \in MDA(H_\gamma)$ for $\gamma \in \mathbb{R}$.

- Choose a threshold u (high enough, by means of suitable statistical approaches) and compute

$$N_u := |\{i \in \{1, 2, \ldots, n\} : X_i > u\}|$$

POT: estimators for the tail and the quantile of the excess distribution

Let $X_1, \ldots, X_n$ i.i.d. r.v. with distribution function $F \in MDA(H_\gamma)$ for $\gamma \in \mathbb{R}$.

- Choose a threshold u (high enough, by means of suitable statistical approaches) and compute

$$N_u := |\{i \in \{1, 2, \ldots, n\} : X_i > u\}|$$

- Let $Y_1, Y_2, \ldots, Y_{N_u}$ be the exceedances. Determine $\hat{\beta}$ and $\hat{\gamma}$, such that the following holds:

$$\bar{F}_u(y) \approx \bar{G}_{\hat{\gamma}, 0, \widehat{\beta(u)}}(y),$$

  where $\bar{F}_u(y) = P(X - u > y | X > u)$.

POT: estimators for the tail and the quantile of the excess distribution

Let $X_1, \ldots, X_n$ i.i.d. r.v. with distribution function $F \in MDA(H_\gamma)$ for $\gamma \in \mathbb{R}$.

- Choose a threshold u (high enough, by means of suitable statistical approaches) and compute

$$N_u := |\{i \in \{1,2,\ldots,n\}: X_i > u\}|$$

- Let $Y_1, Y_2, \ldots, Y_{N_u}$ be the exceedances. Determine $\hat{\beta}$ and $\hat{\gamma}$, such that the following holds:

$$\bar{F}_u(y) \approx \bar{G}_{\hat{\gamma},0,\widehat{\beta(u)}}(y),$$

where $\bar{F}_u(y) = P(X-u > y | X > u)$.

- Use N_u and $\bar{F}_u \approx \bar{G}_{\hat{\gamma},0,\widehat{\beta(u)}}$ to obtain estimators for the tail and the quantile of F

$$\bar{F}\widehat{(u+y)} = \frac{N_u}{n}\left(1+\hat{\gamma}\frac{y}{\hat{\beta}}\right)^{-1/\hat{\gamma}} \text{ and } \hat{q}_p = u + \frac{\hat{\beta}}{\hat{\gamma}}\left(\left(\frac{n}{N_u}(1-p)\right)^{-\hat{\gamma}} - 1\right)$$

POT contd.: How to choose a suitable threshold u?

POT contd.: How to choose a suitable threshold u?

- Is u too large, then there are only a few observed exceedances and not enough data to estimate $\hat{\beta}$ und $\hat{\gamma}$.

POT contd.: How to choose a suitable threshold u?

- Is u too large, then there are only a few observed exceedances and not enough data to estimate $\hat{\beta}$ und $\hat{\gamma}$.
- Is u too small, then the approximation $\bar{F}_u(y) \approx \bar{G}_{\hat{\gamma},0,\hat{\beta}(u)}(y)$ is not good.

POT contd.: How to choose a suitable threshold u?

- Is u too large, then there are only a few observed exceedances and not enough data to estimate $\hat{\beta}$ und $\hat{\gamma}$.
- Is u too small, then the approximation $\bar{F}_u(y) \approx \bar{G}_{\hat{\gamma},0,\hat{\beta}(u)}(y)$ is not good.

Basic idea: inspect the plot of the *empirical mean excess function* and choose a threshold u_0, such that the empirical mean excess function is approximately linear for $u > u_0$.

POT contd.: How to choose a suitable threshold u?

- Is u too large, then there are only a few observed exceedances and not enough data to estimate $\hat{\beta}$ und $\hat{\gamma}$.
- Is u too small, then the approximation $\bar{F}_u(y) \approx \bar{G}_{\hat{\gamma},0,\hat{\beta}(u)}(y)$ is not good.

Basic idea: inspect the plot of the *empirical mean excess function* and choose a threshold u_0, such that the empirical mean excess function is approximately linear for $u > u_0$.

The justification :

- $e_F(u) = \int_0^{x_F - u} t dF_u(t) \approx \int_0^{x_F - u} t dG_{\gamma,0,\beta(u)}(t) = E(G_{\gamma,0,\beta(u)}) = \frac{\beta(u)}{1-\gamma}$, if $F_u(t) \approx G_{\gamma,0,\beta(u)}(t)$.

POT contd.: How to choose a suitable threshold u?

- Is u too large, then there are only a few observed exceedances and not enough data to estimate $\hat{\beta}$ und $\hat{\gamma}$.
- Is u too small, then the approximation $\bar{F}_u(y) \approx \bar{G}_{\hat{\gamma},0,\hat{\beta}(u)}(y)$ is not good.

Basic idea: inspect the plot of the *empirical mean excess function* and choose a threshold u_0, such that the empirical mean excess function is approximately linear for $u > u_0$.

The justification :

- $e_F(u) = \int_0^{x_F - u} t dF_u(t) \approx \int_0^{x_F - u} t dG_{\gamma,0,\beta(u)}(t) = E(G_{\gamma,0,\beta(u)}) = \frac{\beta(u)}{1-\gamma}$, if $F_u(t) \approx G_{\gamma,0,\beta(u)}(t)$.
- If $\bar{F}_u(x) \approx \bar{G}_{\gamma,0,\beta(u)}(x)$ then $\forall v \geq u$ the approximation $\bar{F}_v(x) \approx \bar{G}_{\gamma,0,\beta(u)+\gamma(v-u)}(x)$ holds.

POT (contd.): How to choose a suitable threshold u?

POT (contd.): How to choose a suitable threshold u?

Definition: The empirical mean excess function:
Let $x_1, x_2, \ldots, x_n$ be a sample of i.i.d r.v. Let $N_u = |\{i\colon 1 \leq i \leq n, x_i > u\}|$ be the number of the sample points which exceed u. The empirical mean excess function $e_n(u)$ is defined as:

$$e_n(u) = \frac{1}{N_u} \sum_{i=1}^{n} (x_i - u) I_{\{x_i > u\}}.$$

POT (contd.): How to choose a suitable threshold u?

Definition: The empirical mean excess function:
Let $x_1, x_2, \ldots, x_n$ be a sample of i.i.d r.v. Let $N_u = |\{i \colon 1 \le i \le n, x_i > u\}|$ be the number of the sample points which exceed u. The empirical mean excess function $e_n(u)$ is defined as:

$$e_n(u) = \frac{1}{N_u} \sum_{i=1}^{n} (x_i - u) I_{\{x_i > u\}}.$$

Consider the plot of the (interpolation of the) empirical mean excess function: $(x_{k,n}, e_n(x_{k,n}))$, $k = 1, 2, \ldots, n-1$. If this plot is approximately linear around some $x_{k,n}$, then $u := x_{k,n}$ might be a good choice for the threshold value.

POT (contd.): Estimation of the paramters γ and β

POT (contd.): Estimation of the paramters γ and β

Let u be a given threshold and let $Y_1, Y_2, \ldots, Y_{N_u}$ be the observed data from the sample which exceed u.

POT (contd.): Estimation of the paramters γ and β

Let u be a given threshold and let $Y_1, Y_2, \ldots, Y_{N_u}$ be the observed data from the sample which exceed u.

The likelihood function $L(\gamma, \beta, Y_1, \ldots, Y_{N_u})$ is the conditional probability that $\bar{F}_u(y) \approx \bar{G}_{\gamma,0,\beta}(y)$ under the condition that the observed exceedances are $Y_1, Y_2, \ldots, Y_{N_u}$.

POT (contd.): Estimation of the paramters γ and β

Let u be a given threshold and let $Y_1, Y_2, \ldots, Y_{N_u}$ be the observed data from the sample which exceed u.

The likelihood function $L(\gamma, \beta, Y_1, \ldots, Y_{N_u})$ is the conditional probability that $\bar{F}_u(y) \approx \bar{G}_{\gamma,0,\beta}(y)$ under the condition that the observed exceedances are $Y_1, Y_2, \ldots, Y_{N_u}$.

The following holds:

$$\ln L(\gamma, \beta, Y_1, \ldots, Y_{N_u}) = -N_u \ln \beta - \left(\frac{1}{\gamma} + 1\right) \sum_{i=1}^{N_u} \ln \left(1 + \frac{\gamma}{\beta} Y_i\right)$$

where $Y_i \geq 0$ for $\gamma > 0$ and $0 \leq Y_i \leq -\beta/\gamma$ for $\gamma < 0$.

POT (contd.): Estimation of the paramters γ and β

Let u be a given threshold and let $Y_1, Y_2, \ldots, Y_{N_u}$ be the observed data from the sample which exceed u.

The likelihood function $L(\gamma, \beta, Y_1, \ldots, Y_{N_u})$ is the conditional probability that $\bar{F}_u(y) \approx \bar{G}_{\gamma,0,\beta}(y)$ under the condition that the observed exceedances are $Y_1, Y_2, \ldots, Y_{N_u}$.

The following holds:

$$\ln L(\gamma, \beta, Y_1, \ldots, Y_{N_u}) = -N_u \ln \beta - \left(\frac{1}{\gamma} + 1\right) \sum_{i=1}^{N_u} \ln \left(1 + \frac{\gamma}{\beta} Y_i\right)$$

where $Y_i \geq 0$ for $\gamma > 0$ and $0 \leq Y_i \leq -\beta/\gamma$ for $\gamma < 0$.
(see Daley, Veve-Jones (2003) and Coles (2001))

POT contd.

POT contd.

The maximizers $\hat{\gamma}$ and $\hat{\beta}$ of the log-likelihood function are used as estimators for γ and β (ML-estimators)

POT contd.

The maximizers $\hat{\gamma}$ and $\hat{\beta}$ of the log-likelihood function are used as estimators for γ and β (ML-estimators)

The method works well for $\gamma > -1/2$.

POT contd.

The maximizers $\hat{\gamma}$ and $\hat{\beta}$ of the log-likelihood function are used as estimators for γ and β (ML-estimators)

The method works well for $\gamma > -1/2$.

The ML-estimators are in this case normally distributed:

$$(\hat{\gamma} - \gamma, \frac{\hat{\beta}}{\beta} - 1) \sim N(0, \Sigma^{-1}/N_u) \text{ where } \Sigma^{-1} = \begin{pmatrix} 1+\gamma & -1 \\ -1 & 2 \end{pmatrix}.$$

POT contd.

The maximizers $\hat{\gamma}$ and $\hat{\beta}$ of the log-likelihood function are used as estimators for γ and β (ML-estimators)

The method works well for $\gamma > -1/2$.

The ML-estimators are in this case normally distributed:

$$(\hat{\gamma} - \gamma, \frac{\hat{\beta}}{\beta} - 1) \sim N(0, \Sigma^{-1}/N_u) \text{ where } \Sigma^{-1} = \begin{pmatrix} 1+\gamma & -1 \\ -1 & 2 \end{pmatrix}.$$

There is an uncertainty related to the more or less arbitrary choice of the threshold u. It can be reduced by

POT contd.

The maximizers $\hat{\gamma}$ and $\hat{\beta}$ of the log-likelihood function are used as estimators for γ and β (ML-estimators)

The method works well for $\gamma > -1/2$.

The ML-estimators are in this case normally distributed:

$$(\hat{\gamma} - \gamma, \frac{\hat{\beta}}{\beta} - 1) \sim N(0, \Sigma^{-1}/N_u) \text{ where } \Sigma^{-1} = \begin{pmatrix} 1+\gamma & -1 \\ -1 & 2 \end{pmatrix}.$$

There is an uncertainty related to the more or less arbitrary choice of the threshold u. It can be reduced by

- investigating the dependency of the ML-estimator $\hat{\gamma}$ on u.

POT contd.

The maximizers $\hat{\gamma}$ and $\hat{\beta}$ of the log-likelihood function are used as estimators for γ and β (ML-estimators)

The method works well for $\gamma > -1/2$.

The ML-estimators are in this case normally distributed:

$$(\hat{\gamma} - \gamma, \frac{\hat{\beta}}{\beta} - 1) \sim N(0, \Sigma^{-1}/N_u) \text{ where } \Sigma^{-1} = \begin{pmatrix} 1+\gamma & -1 \\ -1 & 2 \end{pmatrix}.$$

There is an uncertainty related to the more or less arbitrary choice of the threshold u. It can be reduced by

- investigating the dependency of the ML-estimator $\hat{\gamma}$ on u.
- visualizing and inspecting the estimated tail distribution

$$\hat{\bar{F}}(u+y) = \frac{N_u}{n}\left(1 + \hat{\gamma}\frac{y}{\hat{\beta}}\right)^{-1/\hat{\gamma}}$$

Estimation of VaR und CVaR by means of POT

Estimation of VaR und CVaR by means of POT

Let x_1, x_2, ..., x_n be a sample of i.i.d. r.v. with an unknown distribution function F. From the POT method we get the following estimators for the tail distribution and the quantile $q_p = VaR_p(F)$ of F

$$\hat{\bar{F}}(u+y) = \frac{N_u}{n}\left(1+\hat{\gamma}\frac{y}{\hat{\beta}}\right)^{-1/\hat{\gamma}} \text{ and } \hat{q}_p = u + \frac{\hat{\beta}}{\hat{\gamma}}\left(\left(\frac{n}{N_u}(1-p)\right)^{-\hat{\gamma}} - 1\right)$$

Estimation of VaR und CVaR by means of POT

Let $x_1, x_2, \ldots, x_n$ be a sample of i.i.d. r.v. with an unknown distribution function F. From the POT method we get the following estimators for the tail distribution and the quantile $q_p = VaR_p(F)$ of F

$$\hat{\bar{F}}(u+y) = \frac{N_u}{n}\left(1+\hat{\gamma}\frac{y}{\hat{\beta}}\right)^{-1/\hat{\gamma}} \text{ and } \hat{q}_p = u + \frac{\hat{\beta}}{\hat{\gamma}}\left(\left(\frac{n}{N_u}(1-p)\right)^{-\hat{\gamma}} - 1\right)$$

For $\hat{\gamma} \in (0,1)$ we get the following estimator for CVaR:

$$\widehat{CVaR_p}(F) = \hat{q}_p + \frac{\hat{\beta} + \hat{\gamma}(\hat{q}_p - u)}{1-\hat{\gamma}}$$

Estimation of VaR und CVaR by means of POT

Let $x_1, x_2, \ldots, x_n$ be a sample of i.i.d. r.v. with an unknown distribution function F. From the POT method we get the following estimators for the tail distribution and the quantile $q_p = VaR_p(F)$ of F

$$\hat{\bar{F}}(u+y) = \frac{N_u}{n}\left(1+\hat{\gamma}\frac{y}{\hat{\beta}}\right)^{-1/\hat{\gamma}} \text{ and } \hat{q}_p = u + \frac{\hat{\beta}}{\hat{\gamma}}\left(\left(\frac{n}{N_u}(1-p)\right)^{-\hat{\gamma}} - 1\right)$$

For $\hat{\gamma} \in (0,1)$ we get the following estimator for CVaR:

$$\widehat{CVaR_p}(F) = \hat{q}_p + \frac{\hat{\beta} + \hat{\gamma}(\hat{q}_p - u)}{1-\hat{\gamma}}$$

The proof is done in three steps:

Estimation of VaR und CVaR by means of POT (contd.)

(1) For any distribution F and any $p \in (0,1)$ the equality $CVaR_p(F) = q_p + e_F(q_p)$ holds, where $q_p := VaR_p(F)$.

Estimation of VaR und CVaR by means of POT (contd.)

(1) For any distribution F and any $p \in (0,1)$ the equality $CVaR_p(F) = q_p + e_F(q_p)$ holds, where $q_p := VaR_p(F)$.

(2) Let X be a r.v. with $X \sim GPD_{\gamma,0,\beta}$ and $\gamma \notin \{0,1\}$. Then $e_X(q_p) = \frac{\beta+\gamma q_p}{1-\gamma}$, where $q_p := VaR_p(X)$. Thus

$$CVaR_p(X) = q_p + \frac{\beta + \gamma q_p}{1-\gamma}\,.$$

Estimation of VaR und CVaR by means of POT (contd.)

(1) For any distribution F and any $p \in (0,1)$ the equality $CVaR_p(F) = q_p + e_F(q_p)$ holds, where $q_p := VaR_p(F)$.

(2) Let X be a r.v. with $X \sim GPD_{\gamma,0,\beta}$ and $\gamma \notin \{0,1\}$. Then $e_X(q_p) = \frac{\beta+\gamma q_p}{1-\gamma}$, where $q_p := VaR_p(X)$. Thus

$$CVaR_p(X) = q_p + \frac{\beta + \gamma q_p}{1-\gamma}\,.$$

(3) Let X be a r.v. with $X \sim F$. The tail distribution $\bar{F}(x)$ is approximated by $\bar{F}(u)\bar{G}_{\gamma,0,\beta}(x-u)$.

Estimation of VaR und CVaR by means of POT (contd.)

(1) For any distribution F and any $p \in (0,1)$ the equality $CVaR_p(F) = q_p + e_F(q_p)$ holds, where $q_p := VaR_p(F)$.

(2) Let X be a r.v. with $X \sim GPD_{\gamma,0,\beta}$ and $\gamma \notin \{0,1\}$. Then $e_X(q_p) = \frac{\beta+\gamma q_p}{1-\gamma}$, where $q_p := VaR_p(X)$. Thus

$$CVaR_p(X) = q_p + \frac{\beta + \gamma q_p}{1-\gamma}\,.$$

(3) Let X be a r.v. with $X \sim F$. The tail distribution $\bar{F}(x)$ is approximated by $\bar{F}(u)\bar{G}_{\gamma,0,\beta}(x-u)$.

This implies $F \approx \tilde{F}$ with $\tilde{F} := 1 - \bar{F}(u)\bar{G}_{\gamma,0,\beta}(x-u)$.

Estimation of VaR und CVaR by means of POT (contd.)

(1) For any distribution F and any $p \in (0,1)$ the equality $CVaR_p(F) = q_p + e_F(q_p)$ holds, where $q_p := VaR_p(F)$.

(2) Let X be a r.v. with $X \sim GPD_{\gamma,0,\beta}$ and $\gamma \notin \{0,1\}$. Then $e_X(q_p) = \frac{\beta+\gamma q_p}{1-\gamma}$, where $q_p := VaR_p(X)$. Thus

$$CVaR_p(X) = q_p + \frac{\beta + \gamma q_p}{1-\gamma}\,.$$

(3) Let X be a r.v. with $X \sim F$. The tail distribution $\bar{F}(x)$ is approximated by $\bar{F}(u)\bar{G}_{\gamma,0,\beta}(x-u)$.

This implies $F \approx \tilde{F}$ with $\tilde{F} := 1 - \bar{F}(u)\bar{G}_{\gamma,0,\beta}(x-u)$.

The CVaR of the approximation $\tilde{F}$ is given as follows for $q_p > u$:

$$CVaR_p(\tilde{F}) = \hat{q}_p + \frac{\beta + \gamma(\hat{q}_p - u)}{1-\gamma}$$